

0590



ENTERED

RAW SEQUENCE LISTING

PATENT APPLICATION: US/09/966,459A

DATE: 02/26/2002

TIME: 09:40:03

Input Set : A:\30534117.app

Output Set: N:\CRF3\02262002\I966459A.raw

3 <110> APPLICANT: FEDER, J.N.
 4 MINTIER, G.
 5 RAMANATHAN, C.S.
 6 HAWKEN, D.R.
 7 CACACE, A.
 8 BARBER, L.
 9 KORNACKER, M.G.
 11 <120> TITLE OF INVENTION: A NOVEL HUMAN G-PROTEIN COUPLED RECEPTOR, HGPRBMY4,
 12 EXPRESSED HIGHLY IN PROSTATE, COLON, AND LUNG
 14 <130> FILE REFERENCE: D0039NP
 16 <140> CURRENT APPLICATION NUMBER: 09/966,459A
 17 <141> CURRENT FILING DATE: 2001-09-26
 19 <150> PRIOR APPLICATION NUMBER: 60/235,833
 20 <151> PRIOR FILING DATE: 2000-09-27
 22 <150> PRIOR APPLICATION NUMBER: 60/261,776
 23 <151> PRIOR FILING DATE: 2001-01-16
 25 <150> PRIOR APPLICATION NUMBER: 60/305,351
 26 <151> PRIOR FILING DATE: 2001-07-13
 28 <150> PRIOR APPLICATION NUMBER: 60/313,202
 29 <151> PRIOR FILING DATE: 2001-08-17
 31 <160> NUMBER OF SEQ ID NOS: 60
 33 <170> SOFTWARE: PatentIn Ver. 2.1
 35 <210> SEQ ID NO: 1
 36 <211> LENGTH: 957
 37 <212> TYPE: DNA
 38 <213> ORGANISM: Homo sapiens
 40 <400> SEQUENCE: 1
 41 atgatggtgg atcccaatgg caatgaatcc agtgctacat acttcatcct aataggcctc 60
 42 cctggttttag aagaggctca gttctggttg gccttcccat tgtgctccct ctaccttatt 120
 43 gctgtgctag gtaacttgac aatcatctac attgtgcgga ctgagcacag cctgcatgag 180
 44 cccatgtata tatttctttg catgctttca ggcattgaca tctcatctc cacctcatcc 240
 45 atgccccaaa tgctggccat cttctggttc aattccacta ccatccagtt tgatgcttgt 300
 46 ctgctacaga tgtttgccat ccaactcctta tctggcatgg aatccacagt gctgctggcc 360
 47 atggcttttg accgctatgt ggccatctgt caccactgc gccatgccac agtacttacg 420
 48 ttgcctcgtg tcaccaaatt tgggtgtggct gctgtggtgc ggggggctgc actgatggca 480
 49 ccccttccctg tcttcatcaa gcagctgccc ttctgcoct ccaatatcct ttccattcc 540
 50 tactgcctac accaagatgt catgaagctg gcctgtgatg atatccgggt caatgtcgtc 600
 51 tatggcctta tegtcatcat ctccgccatt ggcctggact cacttctcat ctcttctca 660
 52 tatctgctta ttcttaagac tgtgttgggc ttgacacgtg aagcccaggc caaggcattt 720
 53 ggcacttgcg tctctcatgt gtgtgctgtg ttcatattct atgtaccttt cattggattg 780
 54 tccatggtgc atcgctttag caagcggcgt gactctccgc tgcccgctcat cttggccaat 840
 55 atctatctgc tggttcctcc tgtgtcctaac ccaattgtct atggagtga gacaaaggag 900
 56 attcgacagc gcctccttcg acttttccat gtggccacac acgcttcaga gccctag 957

RAW SEQUENCE LISTING

DATE: 02/26/2002

PATENT APPLICATION: US/09/966,459A

TIME: 09:40:03

Input Set : A:\30534117.app

Output Set: N:\CRF3\02262002\I966459A.raw

```

59 <210> SEQ ID NO: 2
60 <211> LENGTH: 318
61 <212> TYPE: PRT
62 <213> ORGANISM: Homo sapiens
64 <400> SEQUENCE: 2
65 Met Met Val Asp Pro Asn Gly Asn Glu Ser Ser Ala Thr Tyr Phe Ile
66   1           5           10           15
68 Leu Ile Gly Leu Pro Gly Leu Glu Glu Ala Gln Phe Trp Leu Ala Phe
69           20           25           30
71 Pro Leu Cys Ser Leu Tyr Leu Ile Ala Val Leu Gly Asn Leu Thr Ile
72           35           40           45
74 Ile Tyr Ile Val Arg Thr Glu His Ser Leu His Glu Pro Met Tyr Ile
75           50           55           60
77 Phe Leu Cys Met Leu Ser Gly Ile Asp Ile Leu Ile Ser Thr Ser Ser
78   65           70           75           80
80 Met Pro Lys Met Leu Ala Ile Phe Trp Phe Asn Ser Thr Thr Ile Gln
81           85           90           95
83 Phe Asp Ala Cys Leu Leu Gln Met Phe Ala Ile His Ser Leu Ser Gly
84           100          105          110
86 Met Glu Ser Thr Val Leu Leu Ala Met Ala Phe Asp Arg Tyr Val Ala
87           115          120          125
89 Ile Cys His Pro Leu Arg His Ala Thr Val Leu Thr Leu Pro Arg Val
90           130          135          140
92 Thr Lys Ile Gly Val Ala Ala Val Val Arg Gly Ala Ala Leu Met Ala
93 145          150          155          160
95 Pro Leu Pro Val Phe Ile Lys Gln Leu Pro Phe Cys Arg Ser Asn Ile
96           165          170          175
98 Leu Ser His Ser Tyr Cys Leu His Gln Asp Val Met Lys Leu Ala Cys
99           180          185          190
101 Asp Asp Ile Arg Val Asn Val Val Tyr Gly Leu Ile Val Ile Ile Ser
102           195          200          205
104 Ala Ile Gly Leu Asp Ser Leu Leu Ile Ser Phe Ser Tyr Leu Leu Ile
105           210          215          220
107 Leu Lys Thr Val Leu Gly Leu Thr Arg Glu Ala Gln Ala Lys Ala Phe
108 225          230          235          240
110 Gly Thr Cys Val Ser His Val Cys Ala Val Phe Ile Phe Tyr Val Pro
111           245          250          255
113 Phe Ile Gly Leu Ser Met Val His Arg Phe Ser Lys Arg Arg Asp Ser
114           260          265          270
116 Pro Leu Pro Val Ile Leu Ala Asn Ile Tyr Leu Leu Val Pro Pro Val
117           275          280          285
119 Leu Asn Pro Ile Val Tyr Gly Val Lys Thr Lys Glu Ile Arg Gln Arg
120           290          295          300
122 Ile Leu Arg Leu Phe His Val Ala Thr His Ala Ser Glu Pro
123 305          310          315
126 <210> SEQ ID NO: 3
127 <211> LENGTH: 1381
128 <212> TYPE: DNA
129 <213> ORGANISM: Homo sapiens

```

RAW SEQUENCE LISTING

DATE: 02/26/2002

PATENT APPLICATION: US/09/966,459A

TIME: 09:40:03

Input Set : A:\30534117.app

Output Set: N:\CRF3\02262002\I966459A.raw

131 <400> SEQUENCE: 3

```

132 ccacgcgtcc gctctgccct gaatccagga tagaccagga caacaagatg agtgggctaac 60
133 tgtaggatgg tgtccatctg tgcctaggg gaggagtagc atcaaaggag aagcaagaac 120
134 tgagaactgt ttggggcact gaagaagtag gactaaggaa gagttagggg gttagtacaa 180
135 atctgaggcc tggttttctg gaaagagacc agagactgac cttattgcat gtcatacaac 240
136 atgcttgctt agagaccctt aatttatattt cttctcttac tctttctgag gaagcatgag 300
137 ccacaccctc agttagtttt gtataatctt aggccttgat agaataaat cttagtcttg 360
138 aaggctttaa aggggaagaa atagctgtct gtgttagtgg tgtgtcagtc agcaggagaa 420
139 cctgctaggg gtggaaggag gagggtagga gtatagccta gacctgagt agataccccg 480
140 ctccaccttg aaagtctctt actggacctc ttatgatgga gttaatacct cctgtttcct 540
141 ctattccaga ttgttttcag ttccagaag gcaaaactga catctcccag gagtccaagt 600
142 aggagattag ggctcccggt ccctatctac tcagtgttag ccttgggcta gagagaggaa 660
143 attcctgctt agaggggaaa atctgcagga cttcgttacc actttcactt tggcagagga 720
144 aggaggtcag ggttggaagg ggaagttagt ctagaatta aaacatagaa ttctgtctac 780
145 aggtggtgga gaggctttct gaaagtgtt ctgggttgag gctgtcacct agattttata 840
146 ttagagttta agtggtccaa aaaattaaga agcagggaag agaaaagaga acaatttcag 900
147 aagcagacga aaggaacagt aataggaaga tctagcaagg atgtggtggg gcagtttcag 960
148 tgtgagatgc catggacagg aaaatggcag catatgtgtg tgtgtgtgtg tgtgtgtgtg 1020
149 tccatgagac agagagacat aaataactaa ataaaaaggc atatcacaaa gaggggctcc 1080
150 tgcttcagct tgagtccttg atgcaaagac atgtggactg ggatcctagc aacctatctg 1140
151 cagccaagga catgacgtta gacgcccaca gaaaaggaaa attggtcaaa cataggaaga 1200
152 gcaactcaagt gccagctaca gtgaatgaca aatacccacc acaagcaca gctctacatt 1260
153 cacaaaaact tggaaaacac aagttcatag actgggcaac cctgagtagt ggagagatca 1320
154 ccagccatgt ttcaggttgt accctctacc tgccctggtg tggtcacagt tcagcttctt 1380
155 c 1381

```

158 <210> SEQ ID NO: 4

159 <211> LENGTH: 2034

160 <212> TYPE: DNA

161 <213> ORGANISM: Homo sapiens

163 <400> SEQUENCE: 4

```

164 gtgtcagtg tcaaaacttct ttccatttca ggtcctctg attcagattt taatgttaac 60
165 attttggaa acagtattca gaaaaaaaat ttcttaata aaaatacaac tcagatcctt 120
166 caaatatgaa actggttggg gaatctccat tttttcaata ttattttott ctttgttttc 180
167 ttgtacata taattattaa taccctgact aggttgtggt tggagggtta ttacttttca 240
168 ttttaccatg cagtccaaat cttaaactgct tctactgatg gtttacagca ttctgagata 300
169 agaatggtac atctagagaa catttgccaa aggcctaagc acggcaaagg aaaataaaca 360
170 cagaatataa taaaatgaga taatctagct taaaactata acttctctt cagaactccc 420
171 aaccacattg gatctcagaa aaatactgtc ttcaaaatga cttctacaga gaagaaataa 480
172 tttttcctct ggacactagc acttaagggg aagattggaa gttaaagcctt gaaaagagta 540
173 catttacctt cgtaaatgaa agttgacaca ctgttctgag agttttcaca gcatatggac 600
174 cctgtttttc ctatttaatt ttcttatcaa ccctttaatt aggcaaagat attattagta 660
175 cctcattgt agccatggga aaattgatgt tcagtgggga tcagtgaatt aaatggggtc 720
176 atacaagtat aaaaattaaa aaaaaaagac ttcatgccc atctcatatg atgtggaaga 780
177 actgttagag agaccaacag ggtagtgggt tagagatttc cagagtctta cattttctag 840
178 aggaggtatt taatttcttc tcaactcttc cagtgttcta tttagggaatt tccctggcaac 900
179 agaactcatg gctttaatcc cactagctat tgcttattgt cctgggtccaa ttgccaatta 960
180 cctgtgtctt ggaagaagtg atttctaggt tcaccattat ggaagattct tattcagaaa 1020
181 gtctgcatag ggcttatagc aagttattta tttttaaaag ttccataggt gattctgata 1080
182 ggcagtgagg ttagggagcc accagttatg atgggaagta tggaatggca ggtcttgaag 1140

```

RAW SEQUENCE LISTING

DATE: 02/26/2002

PATENT APPLICATION: US/09/966,459A

TIME: 09:40:03

Input Set : A:\30534117.app

Output Set: N:\CRF3\02262002\I966459A.raw

```

183 ataacattgg ccttttgagt gtgactcgta gctggaaagt gagggaaatct tcaggacccat 1200
184 gctttatttg gggctttgtg cagtatggaa cagggacttt gagaccagga aagcaatctg 1260
185 acttaggcat gggaatcagg catttttgct tctgaggggc tattaccaag ggtaaatagg 1320
186 tttcatcttc aacaggatat gacaacagtg ttaaccaaga aactcaaatt acaataacta 1380
187 aaacatgtga tcatatatgt ggtaagtttc attttctttt tcaatcctca ggttccctga 1440
188 tatggattcc tataacatgc tttcatcccc ttttgtaatg gatatcatat ttggaaatgc 1500
189 ctatttaata ctgtatttg ctgctggact gtaagcccat gagggcactg tttattattg 1560
190 aatgtcatct ctgttcatca ttgactgctc tttgctcatc attgaatccc ccagcaaagt 1620
191 gcctagaaca taatagtgtc tatgcttgac accggttatt tttcatcaaa cctgattcct 1680
192 tctgtcctga acacatagcc aggcaatttt ccagccttct ttgagttggg tattattaaa 1740
193 ttctggccat tacttccaat gtgagtggaa gtgacatgtg caatttctat acctggctca 1800
194 taaaaccctc ccatgtgcag cctttcatgt tgacattaaa tgtgacttgg gaagctatgt 1860
195 gttacacaga gtaaatcacc agaagcctgg atttctgaaa aaactgtgca gagccaaacc 1920
196 tctgtcattt gcaactccca ctgtatttg tacgaggcag ttggataagt gaaaaataaa 1980
197 gtactattgt gtcaagtcaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaa      2034

```

200 <210> SEQ ID NO: 5

201 <211> LENGTH: 80

202 <212> TYPE: DNA

203 <213> ORGANISM: Artificial Sequence

205 <220> FEATURE:

206 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic

207 oligos

209 <400> SEQUENCE: 5

210 gatccaccat catgaagaag ctgaactgtg accagcacca ggcaggtaga ggctcaaccg 60

211 tatggaagga atgtgtgacc 80

214 <210> SEQ ID NO: 6

215 <211> LENGTH: 20

216 <212> TYPE: DNA

217 <213> ORGANISM: Artificial Sequence

219 <220> FEATURE:

220 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic

221 oligos

223 <400> SEQUENCE: 6

224 actgagcaca gcctgcatga 20

227 <210> SEQ ID NO: 7

228 <211> LENGTH: 25

229 <212> TYPE: DNA

230 <213> ORGANISM: Artificial Sequence

232 <220> FEATURE:

233 <223> OTHER INFORMATION: Description of Artificial Sequence:synthetic

234 oligos

236 <400> SEQUENCE: 7

237 tctgtagcag acaagcatca aactg 25

240 <210> SEQ ID NO: 8

241 <211> LENGTH: 311

242 <212> TYPE: PRT

243 <213> ORGANISM: MOUSE

245 <400> SEQUENCE: 8

246 Met Trp Pro Asn Ser Ser Asp Ala Pro Phe Leu Leu Thr Gly Phe Leu

RAW SEQUENCE LISTING

DATE: 02/26/2002

PATENT APPLICATION: US/09/966,459A

TIME: 09:40:03


Input Set : A:\30534117.app

Output Set: N:\CRF3\02262002\I966459A.raw

```

247   1           5           10           15
249 Gly Leu Glu Met Ile His His Trp Ile Ser Ile Pro Phe Phe Val Ile
250           20           25           30
252 Tyr Phe Ser Ile Ile Val Gly Asn Gly Thr Leu Leu Phe Ile Ile Trp
253           35           40           45
255 Ser Asp His Ser Leu His Glu Pro Met Tyr Tyr Phe Leu Ala Val Leu
256           50           55           60
258 Ala Ser Met Asp Leu Gly Met Thr Leu Thr Thr Met Pro Thr Val Leu
259 65           70           75           80
261 Gly Val Leu Val Leu Asn Gln Arg Glu Ile Val His Gly Ala Cys Phe
262           85           90           95
264 Ile Gln Ser Tyr Phe Ile His Ser Leu Ala Ile Val Glu Ser Gly Val
265           100          105          110
267 Leu Leu Ala Met Ser Tyr Asp Arg Phe Val Ala Ile Cys Thr Pro Leu
268           115          120          125
270 His Tyr Asn Ser Ile Leu Thr Asn Ser Arg Val Met Lys Met Ala Leu
271          130          135          140
273 Gly Ala Leu Leu Arg Gly Phe Val Ser Ile Val Pro Pro Ile Met Pro
274 145          150          155          160
276 Leu Phe Trp Phe Pro Tyr Cys His Ser His Val Leu Ser His Ala Phe
277           165          170          175
279 Cys Leu His Gln Asp Val Met Lys Leu Ala Cys Ala Asp Ile Thr Phe
280           180          185          190
282 Asn Leu Ile Tyr Pro Val Val Leu Val Ala Leu Thr Phe Phe Leu Asp
283           195          200          205
285 Ala Leu Ile Ile Ile Phe Ser Tyr Val Leu Ile Leu Lys Lys Val Met
286          210          215          220
288 Gly Ile Ala Ser Gly Glu Glu Arg Lys Lys Ser Leu Asn Thr Cys Val
289 225          230          235          240
291 Ser His Ile Ser Cys Val Leu Val Phe Tyr Ile Thr Val Ile Gly Leu
292           245          250          255
294 Thr Phe Ile His Arg Phe Gly Lys Asn Ala Pro His Val Val His Ile
295           260          265          270
297 Thr Met Ser Tyr Val Tyr Phe Leu Phe Pro Pro Phe Met Asn Pro Ile
298           275          280          285
300 Ile Tyr Ser Ile Lys Thr Lys Gln Ile Gln Arg Ser Ile Leu Arg Leu
301          290          295          300
303 Leu Ser Lys His Ser Arg Thr
304 305          310
307 <210> SEQ ID NO: 9
308 <211> LENGTH: 307
309 <212> TYPE: PRT
310 <213> ORGANISM: MOUSE
312 <400> SEQUENCE: 9
313 Met Trp Ser Asn Ile Ser Ala Ala Pro Phe Leu Leu Thr Gly Phe Pro
314   1           5           10           15
316 Gly Leu Glu Ala Ala His His Trp Ile Ser Ile Pro Phe Phe Ala Ile
317           20           25           30
319 Tyr Ile Ser Val Leu Leu Gly Asn Gly Thr Leu Leu Tyr Leu Ile Lys

```


 Use of n and / or Xaa has been detected in the Sequence Listing. Review the Sequence Listing to ensure a corresponding explanation is present in the <220> to <223> fields of each sequence using n or Xaa.

VERIFICATION SUMMARY

DATE: 02/26/2002

PATENT APPLICATION: US/09/966,459A

TIME: 09:40:04

Input Set : A:\30534117.app

Output Set: N:\CRF3\02262002\I966459A.raw

L:780 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:15
L:1282 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:1283 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:46
L:1301 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47
L:1302 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:47

ERROR DETECTED**SUGGESTED CORRECTION**SERIAL NUMBER: 09/966,459

ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOF

- 1 **Wrapped Nucleics
Wrapped Aminos** The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."
- 2 **Invalid Line Length** The rules require that a line not exceed 72 characters in length. This includes white spaces.
- 3 **Misaligned Amino
Numbering** The numbering under each 5th amino acid is misaligned. Do not use tab codes between numbers; use space characters, instead.
- 4 **Non-ASCII** The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.
- 5 **Variable Length** Sequence(s) contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.
- 6 **PatentIn 2.0
"bug"** A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) . Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.
- 7 **Skipped Sequences
(OLD RULES)** Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence:
(2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
(i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading)
(xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown)
This sequence is intentionally skipped

Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.
- 8 **Skipped Sequences
(NEW RULES)** Sequence(s) missing. If intentional, please insert the following lines for each skipped sequence.
<210> sequence id number
<400> sequence id number
000
- 9 **Use of n's or Xaa's
(NEW RULES)** Use of n's and/or Xaa's have been detected in the Sequence Listing.
Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present.
In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.
- 10 **Invalid <213>
Response** Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence
- 11 **Use of <220>** Sequence(s) 30 missing the <220> "Feature" and associated numeric identifiers and responses.
Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section.
(See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)
- 12 **PatentIn 2.0
"bug"** Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.
- 13 **Misuse of n** n can only be used to represent a single nucleotide in a nucleic acid sequence. N is not used to represent any value not specifically a nucleotide.